

## Research Article

# BrainNetVis: An Open-Access Tool to Effectively Quantify and Visualize Brain Networks

Eleni G. Christodoulou,<sup>1</sup> Vangelis Sakkalis,<sup>1</sup> Vassilis Tsiaras,<sup>2</sup> and Ioannis G. Tollis<sup>1,2</sup>

<sup>1</sup>*Institute of Computer Science (ICS), Foundation for Research and Technology—Hellas (FORTH), N. Plastira 100, GR-70013 Heraklion, Greece*

<sup>2</sup>*Department of Computer Science, University of Crete, GR-71409 Heraklion, Greece*

Correspondence should be addressed to Eleni G. Christodoulou, echristo@ics.forth.gr

Received 17 September 2010; Revised 25 November 2010; Accepted 31 December 2010

Academic Editor: Sylvain Baillet

Copyright © 2011 Eleni G. Christodoulou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents BrainNetVis, a tool which serves brain network modelling and visualization, by providing both quantitative and qualitative network measures of brain interconnectivity. It emphasizes the needs that led to the creation of this tool by presenting similar works in the field and by describing how our tool contributes to the existing scenery. It also describes the methods used for the calculation of the graph metrics (global network metrics and vertex metrics), which carry the brain network information. To make the methods clear and understandable, we use an exemplar dataset throughout the paper, on which the calculations and the visualizations are performed. This dataset consists of an alcoholic and a control group of subjects.

## 1. Introduction

One of the major issues in neuroscience is to describe how different brain areas communicate with each other during perception, cognition, and action as well as during spontaneous activity in the default or resting state. Mainly two different approaches for capturing and localizing brain activity motifs have been proposed; univariate spectrum based analysis and functional connectivity analysis [1]. Friston [2] defined functional connectivity as the statistical dependence between the activations of distinct and often well-separated neuronal populations.

Network models and graph theory provide a common framework for describing brain functional connectivity [3–5]. The interdependence between brain areas is estimated using multivariate neurophysiological signals (EEG, MEG, ECoG) and/or haemodynamic response images (fMRI). Then, a network is formed by corresponding either brain areas or channels to vertices and by considering an edge between two vertices if and only if the estimated interdependence is above a threshold. Regarding threshold selection, it is important to notice that it is a rather tricky part and there is currently no established way of favouring a specific

threshold value. In practice, a broad range of threshold values is used to characterize the network. However, the authors propose two alternative approaches in selecting a threshold value based either on group statistics between specific graph-theoretic measures of the populations under analysis [6] or utilizing a signal-based technique of selecting the optimal visualization threshold using surrogate (artificially generated ensemble of data aiming at revealing the most significantly coupled brain regions) datasets to correctly identify the most significant correlation patterns [7]. The next step in the analysis, after edge identification, is to measure some networks statistics and characterize the network. Then, using the network characterization, one can draw conclusions on the effect of illnesses or of cognitive loads on functional connectivity [6–11].

In this study, we briefly refer to pairwise (bivariate) and multivariate interdependence measures, as well as linear and nonlinear ones, that have been successfully used as indices of cerebral engagement [12]. This information is important for the correct usage of the tool, especially for nonexpert users, as the application of these measures on the raw EEG data produces the input to our tool. The BrainNetVis tool provides a dynamic snapshot of

the highly complex underlying neural mechanisms by means of graph visualization [13]. BrainNetVis is an open-access multiplatform tool, provided by ICS-FORTH, for graph representation and brain network visualization. Please note that BrainNetVis calculates the following presented metrics on the *synchronization matrices* (adjacency matrices) that the user should calculate in advance! However, the preprocessing section (Section 3.2) briefly presents some widely used techniques to assess functional brain connectivity and form the adjacency matrix.

At this point, we refer to some already existing tools on the field. These tools capture different kinds of EEG information than BrainNetVis and they may be used complementary to it. One of them is EEGLAB [14], which we have been using extensively for better perception of the brain area. EEGLAB is an interactive Matlab toolbox for processing continuous and event-related EEG, MEG, and other electrophysiological data incorporating independent component analysis (ICA), time/frequency analysis, artifact rejection, event-related statistics, and several useful modes of visualization of the averaged and single-trial data. EEGLab offers also dipole localization functions. Some of the metrics that we implement have also been implemented in the Brain Connectivity toolbox (a matlab toolbox) by Rubinov and Sporns [15]. Other related toolboxes include MEA-Tools [16] and ERPWAVELAB [17]. In these toolboxes, however, the measures for quantifying channel interactions are mainly confined to the temporal crosscorrelation [16] and the coherence spectrum [17, 18]. However, more sophisticated interdependence techniques addressing not only linear but also nonlinear synchronization and causality are also available and applied in certain pathologies like Epilepsy [12]. Such measures can act complementary to graph theoretic indices that characterize brain networks as discussed in [19] and can be used as input to BrainNetVis.

The paper is organized as follows. Section 2 presents essential information on the different ways of graph modelling and manipulations, using BrainNetVis. Section 3 refers to the preprocessing needed (Section 3.2), the most commonly used menu calls and the GUI (Section 3.3), and the possible graph visualization options (Section 3.4). Our conclusion is given in Section 4.

## 2. Network Analysis

Before presenting BrainNetVis, it is important to provide here some basic definitions from graph theory.

A *graph*  $G = (V, E)$  defined on a set of *vertices*  $V = \{v_1, \dots, v_n\}$  and *edges*  $E = \{e_1, \dots, e_m\}$ , where each edge  $e \in E$  is an ordered or unordered pair of vertices. An ordered pair  $e = (u, v) \in V \times V$  is called a *directed edge*, while an unordered pair  $e = \{u, v\}$ , where  $u, v \in V$ , is called an *undirected edge*. In case  $u = v$ ,  $e$  is called a *self-loop*. In our study, we consider *simple* graphs, that is, graphs without self-loops. Also the cardinality of  $V$  is denoted by  $n$  (i.e.,  $n = |V|$ ).

A *weighted network*  $G = (V, E, \omega)$  consists of a graph with vertex set  $V$  and edge set  $E$  augmented with an edge

value function  $\omega : E \rightarrow \mathbb{R}$  that assigns to each edge  $e \in E$  a real value  $\omega(e)$ . Every weighted network  $G = (V, E, \omega)$  corresponds to a real  $n \times n$  matrix  $W = (w_{ij})$ ,  $i, j \in \{1, 2, \dots, n\}$ , where  $w_{ij}$  is equal to value  $\omega(e)$  of edge  $e = (v_i, v_j)$  if  $e \in E$ , or to 0 otherwise. If we reserve value 0 to mean the absence of an edge, then the correspondence between  $G$  and  $W$  is one to one. In this work, we consider a subset of weighted networks, which we call *synchronization networks*, where edge values are restricted to interval  $(0, 1]$  and interpreted as strength of dependence between vertices.

In synchronization networks, higher edge values indicate stronger dependencies. To define the length of an edge, we should at least reverse the order of edge values by applying, for example, the inverse function  $g : (0, 1] \rightarrow [1, +\infty)$ , that is,

$$g(x) = \frac{1}{x}. \quad (1)$$

We also propose another function  $g : (0, 1] \rightarrow [1, +\infty)$ , where

$$g(x) = 1 - \log_2(x). \quad (2)$$

These are definitions on how to transform the edge lengths in the case of synchronization networks. Which of the two functions performs better depends on the graph structure and on the metric or the visualization method that uses these functions. When choosing the appropriate formulation, one should consider that the function  $1/x$  tends to  $+\infty$  faster than the function  $1 - \log_2(x)$  when  $x \rightarrow 0^+$ . Therefore, the edges with small values are assigned longer lengths with the  $1/x$  function than those with the  $1 - \log_2(x)$  function.

The length of a path from vertex  $u$  to vertex  $v$  is the sum of the lengths of the edges of the path. The shortest path distance from vertex  $u$  to vertex  $v$  is denoted by  $d_G(u, v)$ . If vertex  $v$  is unreachable from vertex  $u$ , then  $d_G(u, v) = +\infty$ .

## 3. Methods and Results

**3.1. Exemplar Case.** In what follows, we are using the data of a specific use case, consisting of alcoholic and control subjects, in order to provide concrete examples of use of the application. Briefly, the specific study included 30 control subjects and 30 alcoholic subjects. Each subject was fitted with a 61-lead electrode cap (ECI, Electro-Cap International). All scalp electrodes were referred to  $C_z$ . In this experiment, each subject was exposed to pictures of objects chosen from the 1980 Snodgrass and Vanderwart picture set [20]. The stimuli in each trial were randomized (but not repeated) and were presented on a white background for 300 ms at the center of a computer monitor. Their size was approximately 5–10 cm  $\times$  5–10 cm, thus subtending a visual angle of  $0,05^\circ$ – $0,1^\circ$ . Ten trials were shown, with the interval between trials fixed to 3.2 s. The participants were instructed to memorize the pictures in order to be able to identify them later. For each subject and for each trial and frequency band (0.5–4 Hz, 4–8 Hz, 8–13 Hz, 13–30 Hz, 30–45 Hz) the interdependence for each channel pair (there are 61  $(61-1)/2$

channel pairs since the number of active EEG channels is 61) was calculated using the coherence and the RIM methods. The results were stored in  $61 \times 61$  interdependence matrices  $W$  with elements ranging from 0 to 1. The main finding of this study, using BrainNetVis, was that the alcoholic subjects have impaired synchronization of brain activity and loss of lateralization during the rehearsal process as compared to control subjects.

**3.2. Preprocessing.** In order to create a graph, a matrix containing the EEG channel pairwise correlations is required. Thus, one needs to calculate the correlations among all pairs of electrodes and deduce the respective adjacency matrix, called *synchronization matrix*. There exist a number of measures that capture the linear and the nonlinear links between time-series in a frequency band in order to calculate the required correlations (in the EEG analysis context they are called synchronization indices). Three measures have been chosen after an extensive study in linear and nonlinear synchronization measures [12]: the typical magnitude squared coherence method (MSC) [21], a nonlinear bivariate measure for generalized synchronization (RIM) [22] and Partial Directed Coherence (PDC) [23]. The advantage of magnitude squared coherence is that it is well known and widely accepted. The advantage of RIM is that it is able to capture nonlinear patterns available in the signals, whereas PDC can measure causality.

*(1) Magnitude Squared Coherence (MSC)*. MSC (or simply coherence) has been a well-established and traditionally used tool to investigate the linear relation between two signals or EEG channels. Let us suppose that we have two simultaneously measured discrete time series  $x_i$  and  $y_i$ ,  $i = 1 \dots N$ . MSC is the cross-spectral density function  $S_{xy}(f)$ , which is simply derived via the Fourier transform of the crosscorrelation, normalized by their individual autospectral density functions. Hence, MSC is calculated using the Welch's method as

$$\gamma_{xy}(f) = \frac{|\langle S_{xy}(f) \rangle|^2}{|\langle S_{xx}(f) \rangle| |\langle S_{yy}(f) \rangle|}, \quad (3)$$

where  $\langle \cdot \rangle$  indicates window averaging. The estimated MSC for a given frequency  $f$  ranges between 0 (no coupling) and 1 (maximum linear interdependence).

*(2) A Robust Interdependence Measure (RIM)*. Given two scalar time series  $\{x(t)\}_{t \in \mathbb{T}}$  and  $\{y(t)\}_{t \in \mathbb{T}}$  with  $\mathbb{T} = \{1, \dots, N\}$ , which have been measured from dynamical systems  $X$  and  $Y$ , the dynamics of the systems are reconstructed using delay coordinates [24]

$$\mathbf{x}(t) = [x(t), x(t + \tau), \dots, x(t + (m - 1)\tau)]^T \quad (4)$$

and similarly we reconstruct  $\mathbf{y}(t)$  from  $\{y(t)\}_{t \in \mathbb{T}}$ , with an embedding dimension  $m$  and a delay time  $\tau$  for  $n \in \mathbb{T}' = \{1, \dots, N'\}$ , where  $N' = N - (m - 1)\tau$ .

Regarding  $\tau$  and  $m$ , they are parameters of Arnhold's method [25]. Taken's [24] embedding theorems and their

sequels (e.g., [26]) are existence proofs but they do not directly show how to get a suitable time delay  $\tau$  or embedding dimension  $m$  from a finite time series. Empirical and heuristic criteria are employed for selecting  $\tau$  and  $m$ . Usually, a choice of  $\tau$  is the value for which the autocorrelation function first passes through zero, while  $m$  is determined using variations of false nearest neighbour statistics [27–29]. Parameter  $\tau$  can also be calculated using the method of Fraser [30].

Let  $r_{t,j}$  and  $s_{t,j}$ ,  $j = 1, \dots, k$ , denote the time indices of the  $k$  nearest Euclidean neighbors of  $\mathbf{x}(t)$  and  $\mathbf{y}(t)$ , respectively. Temporally correlated neighbors are excluded by means of a Theiler correction:  $|r_{t,j} - t| > m \cdot \tau$  and  $|s_{t,j} - t| > m \cdot \tau$ . For each  $t \in \mathbb{T}'$ , the average square distance of  $\mathbf{y}(t)$  to all remaining points in  $\{\mathbf{y}(j)\}_{j \in \mathbb{T}'}$  is given by

$$R_t(Y) = \frac{1}{N' - 1} \sum_{j=1, j \neq t}^{N'} |\mathbf{y}(t) - \mathbf{y}(j)|^2. \quad (5)$$

For each  $\mathbf{y}_t$ , the X-conditioned mean squared Euclidean distance is defined as

$$R_t^{(k)}\left(\frac{Y}{X}\right) = \frac{1}{k} \sum_{j=1}^k |\mathbf{y}(t) - \mathbf{y}(r_{t,j})|^2. \quad (6)$$

Quiroga et al. [25] defined the dependence measure

$$N\left(\frac{Y}{X}\right) = \frac{1}{N'} \sum_{t=1}^{N'} \frac{R_t(Y) - R_t^{(k)}(Y/X)}{R_t(Y)}. \quad (7)$$

The measure  $N(X/Y)$  is defined in complete analogy, and as interdependence measure between  $X$  and  $Y$ , we use the mean value  $(N(X/Y) + N(Y/X))/2$ .

*(3) Partial Directed Coherence (PDC)*. Let  $\{\mathbf{x}(t)\}_{t \in \mathbb{N}}$  with  $\mathbf{x}(t) = [x_1(t), \dots, x_n(t)]^T$  be a stationary  $n$ -dimensional time series with mean zero. Then, a vector autoregressive model of order  $p$  for  $\mathbf{x}$  is given by

$$\mathbf{x}(t) = \sum_{r=1}^p \mathbf{A}(r)\mathbf{x}(t - r) + \boldsymbol{\varepsilon}(t), \quad (8)$$

where  $\mathbf{A}(r)$  are the  $n \times n$  coefficient matrices of the model and  $\boldsymbol{\varepsilon}(t)$  is a multivariate Gaussian white noise process with covariance matrix  $\boldsymbol{\Sigma}$ . In this model, the coefficients  $A_{ij}(r)$  describe how the present values of  $x_i$  depend linearly on the past values of the components  $x_j$ . In order to provide a frequency domain measure for Granger-causality, Baccala and Sameshima [23] introduced the concept of PDC. This measure is based on the Fourier transform of the coefficient series

$$\bar{\mathbf{A}}(\omega) = I - \sum_{r=1}^p \mathbf{A}(r)e^{-i\omega r}. \quad (9)$$

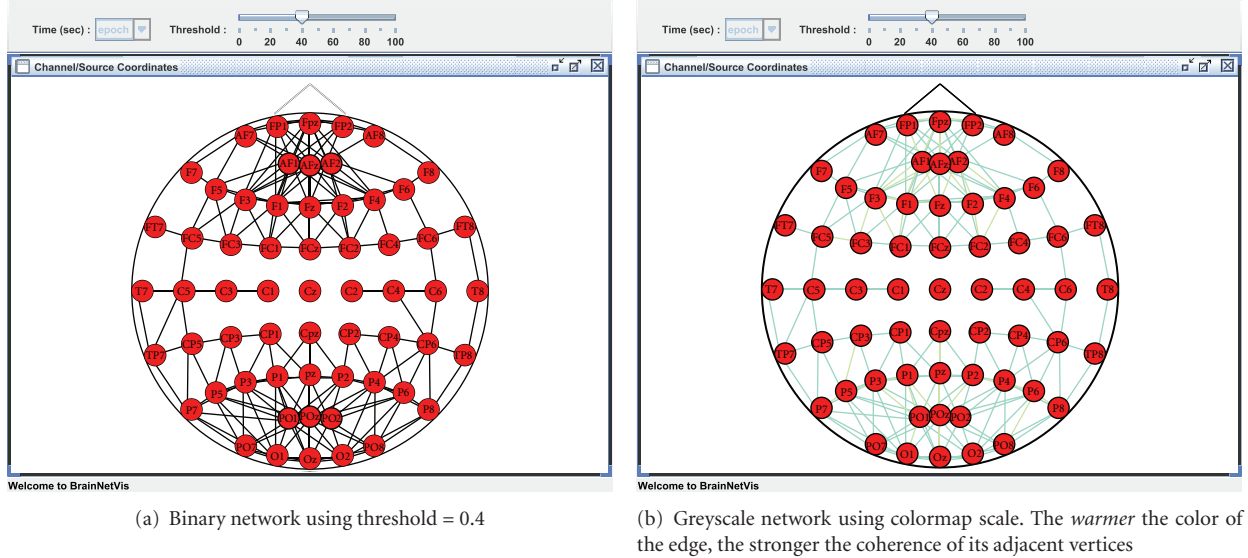


FIGURE 1: Example of weighted networks for a virtual alcoholic patient. Both pictures are produced with the Arnhold's method for broadband activity.

More precisely, the PDC from  $x_j$  to  $x_i$  is defined as

$$\pi_{i-j}(\omega) = \frac{|\overline{A}_{ij}(\omega)|}{\sqrt{\sum_{l=1}^n |\overline{A}_{lj}(\omega)|^2}}. \quad (10)$$

The PDC  $\pi_{i-j}(\omega)$  takes values between 0 and 1 and vanishes for all frequencies  $\omega$  if and only if the coefficients  $A_{ij}(r)$  are zero for all  $r = 1, \dots, p$ .

The synchronization matrix created using one of the above methods serves as input to the BrainNetVis tool thus, it should be calculated separately and a priori. Please note that the presented tool currently implements only graph characterization measures and visualization schemes. It can be used with a variety of inputs in the form of the adjacency matrix. However, we provide the preprocessing section mostly for the interested but not expert user that wishes to investigate how graph analysis may be applied to the neuroscience field. In this sense, even if signal processing techniques are outside of the scope of the tool, we do describe the most widely used methods that provide the input for the further graph analysis. Nevertheless, it is true that most of the methods presented, linear (i.e., PDC) but mostly nonlinear ones (i.e., RIM), assume some kind of *stationarity*. Generally EEG distribution is considered as a multivariate Gaussian process even if the mean and covariance properties generally change from segment to segment. Therefore, strictly speaking, EEG meets quasistationarity because it can be considered stationary only within short intervals. Hence, the user should somehow test the stationarity assumptions prior to using these methods. Hopefully, a novel and prosperous technique capable of decomposing a multivariate time series into its stationary and nonstationary part is known as stationary subspace analysis (SSA) [31] and can be utilized to overcome the implicit stationarity constraints.

**3.2.1. Binary and Greyscale Networks on BrainNetVis.** BrainNetVis provides the option of using either a *binary* or a *greyscale* network by adjusting, respectively, the *Network Metrics Options* under the *View* drop down menu. In our use case, we provided as input to the tool a synchronization matrix describing the brain network of a *virtual alcoholic patient*. This virtual patient has been created by taking the means across the node and edge values over *all 30 alcoholic subjects*. We underline that this subject does not actually exist. We applied a binary network, using *threshold = 0.4* and a greyscale network which we visualized using colormap scale. The edge length transformation function can also be selected under the same menu. We used

$$\ell(e) = \frac{1}{x}. \quad (11)$$

The results are depicted in Figure 1.

**3.2.2. Data Structure.** Two types of files are required for the algorithms that BrainNetVis encapsulates to run properly

- (1) A square synchronization matrix with the data from the EEG study (*required for the algorithms to function*).
- (2) A file containing a matrix of the labels and the coordinates of each electrode. The rows of the table correspond to the electrodes. The first column contains the electrodes' labels, and the other columns contain the coordinates of the electrodes. These will be either 2 columns (for 2D data, respective to  $x$  and  $y$  coordinates) or three columns (for 3D data, respective to  $x$ ,  $y$ , and  $z$  coordinates). (*required for the visualization options*)



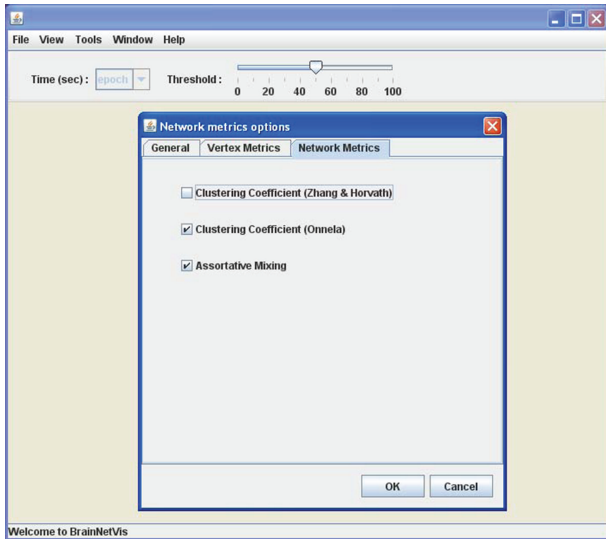


FIGURE 2: A menu screenshot depicting the selection of global network metrics.

3.3. *Menu Calls (GUI)*. The network metrics available in BrainNetVis will be presented here, in a way that follows the tool's structure.

3.3.1. *Global Network Metrics*. Networks are often classified into unifying categories in order to obtain a better understanding of their structure and function. *Network measures* are numbers which capture reduced information for graphs and describe essential properties. Network measures should catch the relevant and needed information, they should differentiate between certain classes of networks and be easily computed in order to be useful in algorithms and applications.

A very important global network metric is *clustering coefficient*. The clustering coefficient has been introduced by Watts and Strogatz [32] in 1998. For a vertex  $v$ , the clustering coefficient  $c(v)$  measures the connectivity of its direct neighborhood. The clustering coefficient  $C(G)$  of a graph is the average of  $c(v)$  taken over all vertices.

In the BrainNetVis application, we implement two different kinds of clustering coefficients, proposed by Zhang and Horvath (the first) and Onnela (the second). Zhang and Horvath proposed a definition which uses only the network values, in the context of gene coexpression networks. On the other hand, Onnela proposed a version of local clustering coefficient based on the concept of subgraph intensity, defined as the geometric average of subgraph edge values. Both metrics are defined in Table 1. It has to be noticed that the Onnela clustering coefficient definition suffers from the drawback that it requires an underlying binary network; if this is not available as a separate set of data, then presumably it must be obtained by discretizing the weighted edges.

The other important global network metric, included in the tool, is *assortative mixing*. This feature captures the similarity between properties of adjacent network vertices. Intuitively, this measure captures the tendency of network

vertices to connect either to vertices with similar degrees (high degrees connected with high degrees and low degrees connected with low degrees) or to vertices that have dissimilar degrees (high degrees connected with low degrees). Newman [33] proposed an interesting measure to quantify the degree of similarity (dissimilarity) between adjacent vertices in a network using assortative mixing, which is given as the correlation between properties of every pairs of adjacent vertices. Each vertex may have assigned a single scalar, such as a centrality measure of the vertex position in a network, or a set of scalar properties. Then, the assortativity coefficient for an undirected graph is defined as the (sample) Pearson product-moment correlation coefficient. The formula of this computation is given in Table 1, and it is written in a symmetrical form. This equation can also be used for directed graphs by simply ignoring the direction of edges.

The value of the assortativity coefficient,  $r$ , lies in the range  $-1 \leq r \leq 1$ , with  $r = 1$  indicating perfect assortativity and  $r = -1$  indicating perfect disassortativity (perfect negative correlation between the properties of the vertices of the edges under consideration). Brain functional networks tend to be assortative [34, 35]. From computational studies, it has been observed that information gets easily transferred through assortative networks as compared to that in disassortative networks [36].

*Global network metrics on BrainNetVis*. BrainNetVis allows the calculation of the mentioned global network metrics by following the *Tools* menu (see Figure 2). Continuing the previous example on an alcoholic patient, we applied the simple *Clustering Coefficient* and the *Assortative Mixing*.

3.3.2. *Vertex Metrics-Centrality Measures*. The above concerned global network metrics. There exists a significant interest in local network properties as well, which concentrates on one node of interest. These properties are very important since at the local scale we can detect which vertices are the most relevant for the organization and functioning of a network. These local measures are commonly named *centrality measures* (or centrality indices) and have proved of great value in analysing the role played by individuals in social networks and in identifying essential proteins, keystone species, and functionally important brain regions.

*Centrality Measures Based on Neighbourhoods*. The simplest and most basic centrality measure is *degree centrality*  $c_D(v)$  of a vertex  $v$ . In practice, this is the number of neighbours of the node of interest. In spite of the simplicity of this concept, degree is the most fundamental network measure and most other centrality measures are linked to it. The definitions of degree centrality, both for directed and for undirected networks are provided in Table 1.

In the case of greyscale networks, instead of using the term *degree centrality*, we use the term *strength centrality*. The formulas for strength centrality are defined correspondingly (Table 1). In BrainNetVis, strength centrality is presented as *normalized degree centrality*. This is accessed when the user chooses the *Normalized Metrics* on the *Tools*  $\Rightarrow$  *Network*

TABLE 1: Network and vertex metrics available in BrainNetVis.

Zhang and Horvath	$c_Z(v) = \sum_{i \neq j \in V \setminus \{v\}} \hat{w}_{vi} \hat{w}_{ij} \hat{w}_{jv} / \sum_{i \neq j \in V \setminus \{v\}} \hat{w}_{vi} \hat{w}_{jv} \Rightarrow$ $c_Z(v) = (1/\max_{i,j}(w_{ij})) \cdot (\sum_{i \neq j \in V \setminus \{v\}} w_{vi} w_{ij} w_{jv} / \sum_{i \neq j \in V \setminus \{v\}} w_{vi} w_{jv})$ <p>The weights have been normalized by <math>\max_{i,j}(w_{ij})</math>.</p> <p>The above definition uses only the network values, in the context of gene coexpression networks.</p>
Onnela	$c_O(v) = (1/(\frac{\deg(v)}{2})) \sum_{i \neq j \in V \setminus \{v\}} (\hat{w}_{vi} \hat{w}_{ij} \hat{w}_{jv})^{1/3} \Rightarrow$ $c_O(v) = (1/\max_{i,j}(w_{ij})) (\frac{\deg(v)}{2}) \sum_{i \neq j \in V \setminus \{v\}} (w_{vi} w_{ij} w_{jv})^{1/3}$ <p>Here, the edge values are normalized by the maximum value in the network,</p> $\hat{w}_{ij} = w_{ij} / \max_{l,k} w_{lk}.$
Assortative mixing	
Symmetrical weighted networks	$r = (4m \sum_{\{u,v\} \in E} \rho(u)\rho(v) - [\sum_{\{u,v\} \in E} (\rho(u) + \rho(v))]^2) / (2m \sum_{\{u,v\} \in E} (\rho(u)^2 + \rho(v)^2) - [\sum_{\{u,v\} \in E} (\rho(u) + \rho(v))]^2)$
Directed weighted networks	$r = (H \sum_{(u,v) \in E} \omega(u,v) \rho(u)\rho(v) - AB) / (\sqrt{H \sum_{(u,v) \in E} \omega(u,v) \rho(u)^2} - A^2 \sqrt{H \sum_{(u,v) \in E} \omega(u,v) \rho(v)^2} - B^2)$ $A = \sum_{(u,v) \in E} \omega(u,v) \rho(u)$ $B = \sum_{(u,v) \in E} \omega(u,v) \rho(v)$ $H = \sum_{e \in E} \omega(e) \text{ is the sum of all values of edges in } E.$
Degree centrality $c_D(v)$ of vertex $v$	
Undirected binary network	Degree $\deg(v)$ of vertex $v$
Directed binary network	In-degree $c_{iD}(v) = \deg^-(v)$ Out-degree $c_{oD}(v) = \deg^+(v)$
Strength centrality $c_S(v)$	
Greyscale symmetric network	Strength $s(v)$ of vertex $v$
Greyscale assymmetric network	In-strength: $c_{iS}(v) = s^-(v)$ Out-strength: $c_{oS}(v) = s^+(v)$
Shortest-path Efficiency	$c_{Ef}(v) = (1/n_{Ef}) \sum_{u \neq v} 1/d_G(v, u)$ , where $n_{Ef} = n - 1$
Shortest-path Betweenness centrality $c_B(v)$ of a vertex $v \in V$	$c_B(v) = (1/n_B) \sum_{s \in V \setminus \{v\}} \sum_{t \in V \setminus \{v,s\}} (\sigma_{st}(v)/\sigma_{st})$ , where $\sigma_{st}$ is the number of shortest $(s, t)$ -paths $\sigma_{st}(v)$ is the number of shortest $(s, t)$ -paths passing through some vertex $v$ other than $s, t$ and $n_B = (n - 1)(n - 2)$ is a normalizing constant.
Bonacich's eigenvector centrality	$\lambda c(v_i) = \sum_{j=1}^n w_{ji} c(v_j)$ <p>In matrix notation with <math>\mathbf{c} = [c(v_1), c(v_2), \dots, c(v_n)]^T</math>, this yields:</p> $\lambda \mathbf{c} = W^T \mathbf{c}.$ <p>This type of equation is well known and solved by the eigenvalues and eigenvectors of <math>W^T</math>.</p> <p>We call the eigenvector <math>\mathbf{s} = [s_1, \dots, s_n]^T</math> of the maximal eigenvalue of <math>\lambda \mathbf{c} = W^T \mathbf{c}</math> principal eigenvector. Then, the eigenvector centrality of node <math>v_i</math> is defined as: <math>c_{EV}(v_i) =  s_i  / \ \mathbf{s}\ _p</math>,</p> <p>where the centrality vector <math>\mathbf{s}</math> is normalized by dividing it by its <math>p</math>-norm</p> $\ \mathbf{s}\ _p = (\sum_{i=1}^n  s_i ^p)^{1/p} \quad 1 \leq p < \infty, \text{ and } \ \mathbf{s}\ _p = \max_{i=1, \dots, n} \{  s_i  \} \quad p = \infty \text{ to produce centrality scores } c(v_i) \leq 1.$
Hubbell's centrality	$\mathbf{c} = \alpha W^T \mathbf{c} + \mathbf{e}$ where $\mathbf{c} = [c(v_1), c(v_2), \dots, c(v_n)]^T$ and $\mathbf{e} = [e_1, e_2, \dots, e_n]^T$ . <p>In order to get meaningful results, <math>\alpha</math> should be chosen according to restriction <math> \alpha  &lt; 1/\lambda_1</math>, where <math>\lambda_1</math> is the maximum value of an eigenvalue of <math>W</math>.</p> <p>This restriction is not mentioned in the literature.</p>
Subgraph centrality of vertex $v_i$	<p>It is given by the <math>i</math>th diagonal entry of the <math>k</math>th power of the adjacency matrix, <math>A</math></p> $c_{SG}(v_i) = \sum_{k=0}^{\infty} \mu_k(i)/k!$ with number of closed walks: $\mu_k(i) = (A^k)_{ii}$ . <p>This measure generalizes to greyscale networks by substituting matrix <math>W</math> for <math>A</math>.</p>

TABLE 1: Continued.

Network entropy	$H(\hat{P}) = -\sum_{i,j} \pi_i \hat{p}_{ij} \log \hat{p}_{ij} = \sum_i \pi_i H_i$ <p>To produce the above equation, we have set a Markov matrix <math>P = [p_{ij}]</math> be the stochastic process which defines the information source and its stationary distribution <math>\pi : \pi P = \pi</math>.</p>
-----------------	---

Metrics Options  $\Rightarrow$  General tab and normalizes the edge values to range from 0 to 1 accordingly.

*Centrality Measures Based on Distances.* Another set of informative measures are the *Centrality Measures Based on Distances*, implying distances that information has to cover in order to be transferred through the network. The first metric that falls in this category is *closeness centrality*. Closeness can be regarded as a measure of how long it will take the information to spread from a given vertex to others in the network. Setting  $G = (V, E)$  as an undirected graph, the shortest path closeness centrality of vertex  $v \in V$  is defined as the inverse of the mean geodesic distance from vertex  $v$  to every other vertex. A serious drawback of this metric is that it can only be used for connected graphs. A new measure, called *shortest path efficiency*, is proposed in Latora and Marchiori [37] and implemented in BrainNetVis application.

For a vertex  $v$ , Latora and Marchiori defined efficiency as

$$ef(v) = \frac{1}{n-1} \sum_{u \neq v} \frac{1}{d_G(v, u)}. \quad (12)$$

The formula for that is provided in Table 1.

Note that (12) can also be used for disconnected graphs. If some vertices  $v$  and  $u$  are not connected, then they do not contribute to  $ef(v)$ . In this case,  $d_G(v, u) = +\infty \Rightarrow 1/d_G(v, u) = 0$ . The global efficiency,  $ef(G)$ , of a graph is the average of  $ef(v)$  taken over all vertices [37]

$$ef(G) = \frac{1}{n} \sum_{v \in V} ef(v) = \frac{1}{n(n-1)} \sum_{v \in V} \sum_{u \neq v} \frac{1}{d_G(v, u)}. \quad (13)$$

In addition to *shortest path efficiency*, we are interested in *shortest-path betweenness centrality*. In this metric, two other nodes, apart from the central vertex  $v$ , are involved. We call these nodes  $s$  and  $t$ , respectively. This metric intuitively refers to the number of shortest paths which connect vertices  $s$  and  $t$  that pass through vertex  $v$ . In the formula provided in Table 1, the relative numbers  $\sigma_{st}(v)/\sigma_{st}$  are interpreted as the extent to which vertex  $v$  controls the communication between vertices  $s$  and  $t$ . A vertex is considered central, if it is between many pairs of other vertices. Shortest-path betweenness centrality can be generalized to greyscale networks where the length of a path is equal to the sum of the lengths of its edges.

*Centrality measures based on Neighborhoods and on Distances in BrainNetVis.* We applied the above types of centrality measures on our synchronization matrix of the alcoholic patient's EEG. Figure 3 depicts the visualization of the individual's brain network using the *Static Visualization Method*. The *Binary Network* using threshold = 0.4 has

been selected. The centrality measures calculated are the *Degree Centrality*, *Shortest Path Efficiency* and *Shortest Path Betweenness Centrality*. They are depicted on the respective table, shown in the same figure. Both the figure and the table with the metrics can be created by the following the *View* menu.

*Spectral Centrality Measures.* Another set of network metrics is based on the calculation of the *eigenvectors* of the adjacency matrix of the network, produced at the preprocessing step. Most of them are calculated by solving a linear equation system. These measures are called *Spectral Centrality Measures*. *Bonacich's eigenvector centrality* is one of them according to which the centrality of each vertex is proportional to the sum of the centralities of the vertices to which it is directly connected. The respective formula is presented in Table 1.

Expanding the simple Bonacich's eigenvector centrality, Hubbell [38] suggested yet another centrality measure based on the solution of a system of linear equations. *Hubbell's centrality* uses an approach based on directed weighted graphs where the weights of the edges may be real numbers. The general assumption of Hubbell's centrality is similar to the idea of Bonacich, but the centrality of a vertex depends both on its connection to other vertices and to exogenous input which sometimes is called boundary conditions. In this case, we include one more input to the equation  $\lambda \mathbf{c} = W^T \mathbf{c}$  which describes Bonacich's eigenvector centrality. The result is shown on Table 1. This formula encapsulates the relative importance of endogenous versus exogenous factors in the determination of centrality.

The next spectral centrality measure, *subgraph centrality*, has been introduced by Estrada et al. [39]. It is calculated as the weighted sum of the number of closed walks in a graph, where longer walks receive lower weight than shorter ones. Very relative to the subgraphs of the network is the number of short walks of length  $k$ , starting and ending on vertex  $v_i$ . This number is symbolized with  $\mu_k(i)$  on Table 1.

Last but not least, a very interesting idea was suggested by Demetrius et al. [40], describing *network entropy*. Evidence has been presented that this quantity is related to the capacity of the network to withstand random changes in the network structure. Network entropy is based on the Kolmogorov-Sinai (KS) entropy, which is a generalization of the Shannon entropy in that it describes the rate at which a stochastic process generates information. In our context, information corresponds to a sequence of vertices visited by an assumed Markov process on the network. Network entropy takes into account the impact of a vertex's removal on the network. This is captured by the product  $\pi_i H_i$  of the respective definition on Table 1. The interested reader could find more detailed information in [41].

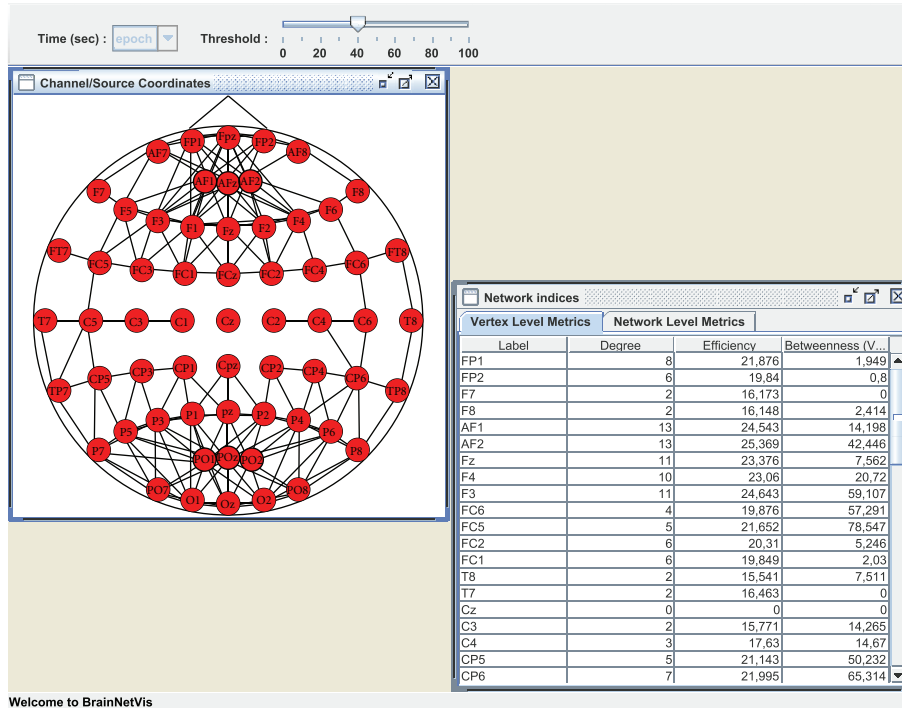


FIGURE 3: Centrality measures for the virtual alcoholic patient based on neighborhoods and on distances in BrainNetVis. The graph has been calculated by the Arnhold's method for broadband activity.

*Spectral Centrality Measures in BrainNetVis.* We applied the above types of centrality measures on our synchronization matrix of the alcoholic patient's EEG. Using links from the *Tools* menu, we calculated the *Bonacich's Eigenvector Centrality*, *Hubbell's Centrality*, *Subgraph Centrality*, and *Network Entrophy*. One can define the type of networks with which he wishes to work (binary or greyscale) and also select the threshold value.

**3.4. Graph Drawing Techniques.** Regarding the way in which the brain is depicted, BrainNetVis tool incorporates three different kinds of visualization as the follows.

**3.4.1. Static Visualization Method.** In this method, in order to visualize the topology of the emerged network, we create a static framework where each electrode is depicted by a node placed in a position similar to the actual electrode's position on the human cortex. Depending on the number of the electrodes of each experiment, an oval shape is outlined (which corresponds to the scalp) and inside this oval shape, a number  $V$  of circles exist that correspond to the electrodes placed on the subjects' head during the experiments.

**3.4.2. Multidimensional Scaling.** Multidimensional Scaling (MDS) is a family of techniques for analysis and visualization of complex data. The "beauty" of MDS is that we can analyze any kind of distance or similarity matrix, in addition to correlation matrices. Objects in a data set are represented as points in a geometric space; distance in this space represents proximity or similarity among objects. In our case, the

objects are the electrodes and the distances among them are respective to their correlation in the synchronization matrix. In general, the goal of the analysis is to detect meaningful underlying connections among the electrodes which reflect the connections among different brain functional regions. In *BrainNetVis*, we incorporated a 2D visualization of the connections among electrodes. At this point, it has to be noticed that the more dimensions we use in order to reproduce the distance matrix, the better the fit of the reproduced is matrix to the observed matrix (i.e., the smaller the stress is). In fact, if we use as many dimensions as there are variables, then we can perfectly reproduce the observed distance matrix. Of course, our goal is to reduce the observed complexity of nature, that is, to explain the distance matrix in terms of fewer underlying dimensions. Some exemplar views of multidimensional scaling are shown in Figure 4

**3.4.3. Force-Based or Force-Directed Algorithms.** These are a class of algorithms for drawing graphs in an aesthetically pleasing way. Their purpose is to position the nodes of a graph in two-dimensional or three-dimensional space so that all the edges are of more or less equal length and there are as few crossing edges as possible. The force-directed algorithms achieve this by assigning forces amongst the set of edges and the set of nodes; the most straightforward method is to assign forces as if the edges were springs (see Hooke's law), and the nodes were electrically charged particles (see Coulomb's law). The entire graph is then simulated as if it were a physical system. The forces between its nodes change the dynamics and the layout of the system which at







- (ii) *Network Metrics*. Following this tab, the user can ask either for the *Vertex level metrics* table, which contains the values of the vertex metrics that interest the user (and which he chooses under the *Tools* drop-down menu), or for the *Network level metrics*, which contains the values of the global network metrics.

*Tools*. This menu contains the following.

- (i) *Display Options*. Following this tab, the user can set up the display of the graphs. He can set his preferences concerning the nodes (size, color, label, font) and/or the edges (size, color, direction, arrow size).
- (ii) *Network Metrics Options*. Three tabs appear in this sub-menu. The first one is named *General* and contains options like if the network is directed or not, binary or not and synchronization network or not. In the latter case, the tool provides an option on the normalization of the edge length. The second tab is named *Vertex Metrics* and contains options for all the vertex metrics described in Section 3.3.2. Finally, the last tab is named *Network Metrics* and contains options for the network metrics described in Section 3.3.1.

*Window*. Here, the user can change the size of the window of the GUI.

## Acknowledgment

The authors wish to thank Dimitris Andreou for the development of the supportive software of the tool's different versions.

## References

- [1] V. Sakkalis, "Applied strategies towards EEG/MEG biomarker identification in clinical and cognitive research," *Biomarkers in Medicine*, vol. 5, no. 1, pp. 93–105, 2011.
- [2] K. J. Friston, "Functional and effective connectivity in neuroimaging: a synthesis," *Human Brain Mapping*, vol. 2, no. 1-2, pp. 56–78, 1994.
- [3] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nature Reviews Neuroscience*, vol. 10, no. 3, pp. 186–198, 2009.
- [4] C. J. Stam and J. C. Reijneveld, "Graph theoretical analysis of complex networks in the brain," *Nonlinear Biomedical Physics*, vol. 1, article 3, 2007.
- [5] F. De Vico Fallani, L. Astolfi, F. Cincotti et al., "Brain network analysis from high-resolution EEG recordings by the application of theoretical graph indexes," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 16, no. 5, pp. 442–452, 2008.
- [6] V. Sakkalis, T. Oikonomou, E. Pachou, I. Tollis, S. Micheloyannis, and M. Zervakis, "Time-significant wavelet coherence for the evaluation of schizophrenic brain activity using a graph theory approach," in *Proceedings of the 28th IEEE-EMBS, Engineering in Medicine and Biology Society (EMBC '06)*, vol. 1, pp. 4265–4268, New York, NY, USA, 2006.
- [7] V. Sakkalis, V. Tsiaras, M. Zervakis, and I. Tollis, "Optimal brain network synchrony visualization: application in an alcoholism paradigm," in *Proceedings of the 29th Annual International Conference of IEEE-EMBS, Engineering in Medicine and Biology Society (EMBC '07)*, pp. 4285–4288, 2007.
- [8] C. J. Stam, B. F. Jones, G. Nolte, M. Breakspear, and P. Scheltens, "Small-world networks and functional connectivity in Alzheimer's disease," *Cerebral Cortex*, vol. 17, no. 1, pp. 92–99, 2007.
- [9] N. Situ, R. Rezaie, A. Papanicolaou, L. Pollonini, U. Patidar, and G. Zouridakis, "Functional connectivity networks in the autistic and healthy brain assessed using granger causality," in *Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2010.
- [10] M. Massimini, F. Ferrarelli, R. Huber, S. K. Esser, H. Singh, and G. Tononi, "Neuroscience: breakdown of cortical effective connectivity during sleep," *Science*, vol. 309, no. 5744, pp. 2228–2232, 2005.
- [11] M. Valencia, M. A. Pastor, M. A. Fernández-Seara, J. Artieda, J. Martinerie, and M. Chavez, "Complex modular structure of large-scale brain networks," *Chaos*, vol. 19, no. 2, Article ID 023119, 2009.
- [12] V. Sakkalis, C. Doru Giurcăneanu, P. Xanthopoulos et al., "Assessment of linear and nonlinear synchronization measures for analyzing EEG in a mild epileptic paradigm," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 4, pp. 433–441, 2009.
- [13] G. Di Battista, P. Eades, R. Tamassia, and I. G. Tollis, *Graph Drawing: Algorithms for the Visualization of Graphs*, Prentice Hall, Upper Saddle River, NJ, USA, 1999.
- [14] <http://sccn.ucsd.edu/eeglab/>.
- [15] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: uses and interpretations," *NeuroImage*, vol. 52, no. 3, pp. 1059–1069, 2010.
- [16] U. Egert, TH. Knott, C. Schwarz et al., "MEA-Tools: an open source toolbox for the analysis of multi-electrode data with MATLAB," *Journal of Neuroscience Methods*, vol. 117, no. 1, pp. 33–42, 2002.
- [17] M. Mørup, L. K. Hansen, and S. M. Arnfred, "ERP-WAVE-LAB: a toolbox for multi-channel analysis of time-frequency transformed event related potentials," *Journal of Neuroscience Methods*, vol. 161, no. 2, pp. 361–368, 2007.
- [18] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of Neuroscience Methods*, vol. 134, no. 1, pp. 9–21, 2004.
- [19] V. Sakkalis, V. Tsiaras, and I. Tollis, "Graph analysis and visualization for brain function characterization using EEG data," *Journal of Healthcare Engineering*, vol. 1, no. 3, pp. 435–460, 2010.
- [20] J. G. Snodgrass and M. Vanderwart, "A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity," *Journal of Experimental Psychology: Human Learning and Memory*, vol. 6, no. 2, pp. 174–215, 1980.
- [21] S. M. Kay, *Modern Spectral Estimation*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1988.
- [22] J. Arnhold, P. Grassberger, K. Lehnertz, and C. E. Elger, "A robust method for detecting interdependences: application to intracranially recorded EEG," *Physica D*, vol. 134, no. 4, pp. 419–430, 1999.
- [23] L. A. Baccalá and K. Sameshima, "Partial directed coherence: a new concept in neural structure determination," *Biological Cybernetics*, vol. 84, no. 6, pp. 463–474, 2001.

- [24] F. Takens, "Detecting strange attractors in turbulence," in *Proceedings of the Dynamical Systems and Turbulence Symposium*, vol. 898 of *Lecture Notes in Mathematics*, pp. 366–381, 1981.
- [25] R. Q. Quiroga, A. Kraskov, T. Kreuz, and P. Grassberger, "Performance of different synchronization measures in real data: a case study on electroencephalographic signals," *Physical Review E*, vol. 65, no. 4, Article ID 041903, 14 pages, 2002.
- [26] T. Sauer, J. A. Yorke, and M. Casdagli, "Embedology," *Journal of Statistical Physics*, vol. 65, no. 3-4, pp. 579–616, 1991.
- [27] M. B. Kennel, R. Brown, and H. D. I. Abarbanel, "Determining embedding dimension for phase-space reconstruction using a geometrical construction," *Physical Review A*, vol. 45, no. 6, pp. 3403–3411, 1992.
- [28] L. Cao, "Practical method for determining the minimum embedding dimension of a scalar time series," *Physica D*, vol. 110, no. 1-2, pp. 43–50, 1997.
- [29] R. Hegger, H. Kantz, and T. Schreiber, "Practical implementation of nonlinear time series methods: the TISEAN package," *Chaos*, vol. 9, no. 2, pp. 413–435, 1999.
- [30] A. M. Fraser and H. L. Swinney, "Independent coordinates for strange attractors from mutual information," *Physical Review A*, vol. 33, no. 2, pp. 1134–1140, 1986.
- [31] P. Von Bünaeu, F. C. Meinecke, F. C. Király, and K. R. Müller, "Finding stationary subspaces in multivariate time series," *Physical Review Letters*, vol. 103, no. 21, Article ID 214101, 2009.
- [32] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.
- [33] M. E. J. Newman, "Assortative mixing in networks," *Physical Review Letters*, vol. 89, no. 20, Article ID 208701, 4 pages, 2002.
- [34] C. H. Park, S. Y. Kim, Y. H. Kim, and K. Kim, "Comparison of the small-world topology between anatomical and functional connectivity in the human brain," *Physica A*, vol. 387, no. 23, pp. 5958–5962, 2008.
- [35] V. M. Eguiluz, D. R. Chialvo, G. A. Cecchi, M. Baliki, and A. V. Apkarian, "Scale-free brain functional networks," *Physical Review Letters*, vol. 94, no. 1, Article ID 018102, 2005.
- [36] R. Xulvi-Brunet and I. M. Sokolov, "Reshuffling scale-free networks: from random to assortative," *Physical Review E*, vol. 70, no. 6, Article ID 066102, 6 pages, 2004.
- [37] V. Latora and M. Marchiori, "Efficient behavior of small-world networks," *Physical Review Letters*, vol. 87, no. 19, Article ID 198701, 4 pages, 2001.
- [38] C. H. Hubbell, "An input-output approach to clique identification," *Sociometry*, vol. 28, pp. 377–399, 1965.
- [39] E. Estrada and J. A. Rodríguez-Velázquez, "Subgraph centrality in complex networks," *Physical Review E*, vol. 71, no. 5, Article ID 056103, pp. 1–9, 2005.
- [40] L. Demetrius, V. M. Gundlach, and G. Ochs, "Complexity and demographic stability in population models," *Theoretical Population Biology*, vol. 65, no. 3, pp. 211–225, 2004.
- [41] V. L. Tsiaras, *Algorithms for the analysis and visualization of biomedical networks*, Ph.D. thesis, Computer Science Department, University of Crete, Heraklion, Greece, 2009.